

Data Access Governance

Mahsa Shabani, Adrian Thorogood and Madeleine Murtagh

19.1 INTRODUCTION

Enabling researchers' access to large volumes of health data collected in both research and healthcare settings can accelerate improvements in clinical practice and public health. Because the source and subject of those data are people, data access governance has been of concern to scientists, ethics and regulatory scholars, policymakers and citizens worldwide. While researchers have long provided colleagues access to data in an ad hoc fashion, many research funders – e.g. US National Institutes of Health, Wellcome, Bill and Melinda Gates Foundation, United Kingdom Research and Innovation, European Research Council – journals,¹ professional societies and associations,² and regulators now systematically promote the deposit of research data in repositories that aim to provide responsible and timely access to data. Data sharing aims to enable meta-analyses and creative (re)uses, reduce duplicative effort in data generation, and improve reproducibility through validation studies, so as to support data-intensive research and thereby improve human health. In many countries, routinely collected healthcare data is also increasingly being made available to researchers. In both research and healthcare contexts, technical and governance strategies for promoting responsible data sharing and access continue to evolve.

The broad sharing of health research data promises many benefits, but it can also involve risks. Health research data can reveal sensitive information about individuals (in legal terms, data subjects) and their relatives, posing risks to privacy and of discrimination and stigmatisation. Broad sharing of health research data can also raise professional concerns for the researchers or organisations who produce data in terms of receiving adequate credit and recognition for their efforts in collecting, curating and analysing data.³ Likewise, commercial research companies may be concerned their data will be appropriated or misused by competitors. Data access governance aims to promote organisational, scientific and societal interests in data re-use, while protecting the rights and interests of the range of stakeholders with an interest in data. Data access governance manages who has access to data, for what purposes, and under what

¹ D. Taichman et al., 'Sharing Clinical Trial Data: A Proposal from the International Committee of Medical Journal Editors', (2016) *Annals of Internal Medicine*, 164(7), 505–506.

² ACMG Board of Directors, 'Laboratory and Clinical Genomic Data Sharing Is Crucial to Improving Genetic Health Care: A Position Statement of the American College of Medical Genetics and Genomics', (2017) *Genetics in Medicine*, 19(7), 721–722.

³ M. Murtagh et al., 'International Data Sharing in Practice: New Technologies Meet Old Governance', (2016) *Biopreservation and Biobanking*, 14(3), 231–240.

conditions. Governance mechanisms include policies, due diligence processes, data access agreements and monitoring. Data access governance is closely linked to the concept of data stewardship, where organisations aim to ensure data are shared widely in the interest of science and society, while also mitigating associated ethical, societal and privacy risks.⁴

In contemporary data-driven science, data access governance often involves Data Access Committees (DACs) as the key institutional setting in which access decisions are made. DACs are diverse and may be composed of individuals with a range of relevant expertise, including familiarity with the scientific area, privacy and security, and research ethics.⁵ As Lowrance notes, ‘...[s]ome DACs are formally constituted and appointed, while some are more casual. Some publish their criteria, decisions and decision rationales, but most don’t. Some directly advise the data custodians, who then make the yes/no (or revise-and-reapply) access decisions. But many DACs make binding decisions’.⁶

Against this backdrop, this chapter examines the topic of data access governance. We discuss the underlying values and goals of data access governance, focusing in particular on the scientific and social implications for open access and data sharing, on the rights and interests of data subjects as well as those of data producers, and on the ethical conduct of data sharing. We contrast the general structural and normative components of open and controlled data access. We then present existing data access arrangements of organisations and repositories that exemplify varying modes of good practice. We argue these models exemplify the tension between promoting open access to databases on the one hand, and, on the other, protecting the rights and interests of the parties involved, including data subjects, researchers, funding organizations and commercial entities. We suggest that principles of transparency, fairness and proportionality in consideration of all stakeholders’ interests and values is key to achieving this balance. We conclude by discussing existing challenges in data access governance, including potential conflicts between various stakeholders’ views and interests, resource issues, (mis) coordination between oversight bodies, and the need for better harmonisation of access policies and procedures.

19.2 GOALS OF DATA ACCESS GOVERNANCE

Key goals of data access governance aim to strike a balance between protecting data subjects and data producers’ rights and interests, while also promoting broad access to data to advance scientific research in the public interest.

19.2.1 *Protecting Data Subject Rights and Interests and Promoting Research Integrity and Ethics*

Data access governance supports research ethics principles for research involving human subjects. Minimising privacy risks to participants, respecting participant autonomy, and holding researchers accountable for the scientific validity and ethical conduct of research through research ethics committee (REC) approval and oversight, are key goals of governance

⁴ The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, ‘Accessing Health and Health-Related Data in Canada’, (Council of Canadian Academies, 2015).

⁵ Murtagh et al., ‘Better Governance, Better Access: Practising Responsible Data Sharing in the METADAC Governance Infrastructure’, (2018) *Human Genomics*, 12(1), 24.

⁶ W. W. Lowrance, *Privacy, Confidentiality, and Health Research* (Cambridge University Press, 2012).

of data access.⁷ These goals are increasingly furthered by engaging communities in the design of governance.

Privacy and security: Data access governance can protect participant privacy in several ways. Data access agreements, which are signed by data custodians and data users, typically include requirements regarding protecting privacy and security. Privacy safeguards include restrictions on unauthorised individual-level linkage of datasets, which may increase the re-identifiability of data, or prohibitions on attempting to re-identify participants. The greater the combinations of individual-level data for any given individual, the more likely re-identification becomes. Privacy rules in access processes are therefore often designed to control the level of individual-level data linkage. Security safeguards may include general or specific requirements to adopt physical, organisational, and technical protections, as well as data breach reporting obligations.

Respect for the provisions of ethical approvals: Data access governance models often aim to ensure users respect high standards of scientific integrity, and meet the ethical requirements related to compatibility of downstream use of data with the original consent obtained from the participants at the time of enrolment to a study and data collection. Where researchers have stated that data will only be used for certain kinds of research – e.g. disease-specific – this condition will inform the review of an access proposal by the relevant oversight bodies, notably DACs. Data access review may be informed by the following questions:⁸ Does the application violate – or potentially violate – any of the ethical permissions granted to the study or any of the consent forms signed by the study participants or their guardians? Does the application run a significant risk of upsetting or alienating study participants or thereby reducing their willingness to remain as active participants in the research? Does the application run a significant risk of bringing disrepute to study, repository or steward and thereby reducing participant trust and willingness to remain as active participants in the research?

Respect for communities and relevant stakeholders: Responding to relevant stakeholders including communities' concerns and seeking to strike a balance between the views of different groups is fundamental to respecting these communities. This may mean championing the rights of less powerful groups and taking steps to seek out their views and actively responding to those views. In the context of data access, stakeholders include study participants and communities who provide the data, study managers and the researchers who develop the data and related resources, researchers who wish to access those data, the funders who support the studies which produce the data and the public who are the ultimate funders as well as beneficiaries of research. Each of these groups has a legitimate and vested interest in the responsible and respectful uses of data and provide a unique perspective on how such governance can be achieved. For example, study participants and community representatives sitting on oversight committees such as DACs can provide a unique insight into what other study participants may view as acceptable uses of data.

19.2.2 Data Producer Rights and Interests

One goal of access controls is to protect the rights and interests of the researchers or institutions generating data. Academically, researchers compete for high-impact publications and, in turn, for academic positions and promotions. Commercially, researchers and research institutions

⁷ M.Aitken et al., 'Consensus Statement on Public Involvement and Engagement with Data-Intensive Health Research,' (2019) *International Journal of Population Data Science*, 4(1).

⁸ Murtagh et al., 'METADAC Governance Infrastructure', 24.

may compete to develop commercial applications from research findings. These considerations are often addressed through publication and commercialisation clauses in data access agreements.

Data access governance may include publication policies that seek to ensure that data producers are appropriately recognised for their contribution to science. Given that publication remains the major currency in academia, there may be a tendency for data producers to request co-authorship as a condition of access. This is discouraged for reasons of scientific freedom and accountability. Having independent DAC members adjudicating access is one remedy to the potential conflicts of interest in such practices. A compromise position is sometimes used whereby the data producer has a right to review manuscripts before publication, or to at least to be informed in advance of forthcoming publications based on (re)analysis of shared datasets. Commercialisation policies aim to ensure that the data producer benefits from, or at least does not have its competitive position harmed by, downstream use of data.

Finally, responsible data access governance requires transparency, fairness and proportionality towards participants and other stakeholders. Transparency can be improved by the publishing of policies and procedures, as well as publication of approved data recipients and plain language summaries or abstracts of approved uses. Moreover, ensuring timely and consistent access review without imposing unnecessary constraints on data access are of salient importance with regard to fairness. Where data governance seeks to achieve competing goals of openness and privacy protection, as well as meeting social and participant expectations of data use, a proportionate balance needs to be struck. Proportionality may call for different types of access controls to be applied to different types of data. Increasingly, there is emphasis that the balance between public benefit and individual risks be evidence-based.⁹

19.3 DATA ACCESS GOVERNANCE: POLICIES, PROCESSES, AGREEMENTS AND OVERSIGHT

The values and goals of data access governance are operationalised through the policies and practices of DACs and various models of data access.

19.3.1 *Controlled Versus Open Access Data*

The nature of data – and the associated ethical, policy and legal issues – largely determines the access model, which can range from open to controlled to closed. Open access models generally make data available to any user, anywhere, over the internet, without financial or technical constraints. The Human Genome Project, for example, which sequenced the entire human genome, shared the sequence data openly. Subsequent publicly-funded projects sequenced more individuals and combined these data with richer social, demographic and clinical data, prompting concerns about the privacy of data subjects. Controlled access models emerged to ensure data could still be shared broadly with qualified and trusted researchers, while also protecting the privacy of data subjects and sometimes also the interests of researchers producing data. In controlled access, access is managed by a REC or increasingly by a specialised DAC,

⁹ M. Shabani et al., 'Who Should Have Access to Genomic Data and How Should They Be Held Accountable? Perspectives of Data Access Committee Members and Experts', (2016) *European Journal of Human Genetics*, 24(12), 1671–1675; P. Burton et al., 'Policies and Strategies to Facilitate Secondary Use of Research Data in the Health Sciences', (2017) *International Journal of Epidemiology*, 46(6), 1729–1733.

which reviews requests for data access. In this regard, DACs often carry out a due diligence review of access requests and may hold deliberations over the scientific, feasibility and ethical aspects of the request. This is in line with the recommendations issued by the Organisation for Economic Co-operation and Development's (OECD) Council on Health Data Governance that review and approval processes should involve an evidence-based assessment and adhere to principles of transparency, objectivity and fairness. In addition, the OECD's recommendations underline the importance of independent multi-disciplinary review with an ultimate aim of risk mitigation for individuals and society.¹⁰

19.3.2 *Data Access Agreements*

One component of both controlled and open access models is the data agreement (termed 'data transfer', 'data access' or 'data use' agreement), which establishes the conditions governing the accessing researcher's use of the data. The terms of data access agreements typically address data subject protections, including prohibition on unauthorised linkage of individual-level data and attempts to re-identify participants, respect for consent-based use conditions and ensuring appropriate security safeguards are in place. The terms may also include protections for the rights and interests of the researchers producing data, such as publication embargoes to allow data producers the first attempt at publication or intellectual property clauses governing ownership of downstream commercialisation. Benefit-sharing clauses are important in countries with emergent research infrastructures. Other clauses may serve multiple stakeholders, such as obligations to only use data for specified purposes. Still other clauses may address the interests of science and society, such as requirements for open access publication, or to share analysis code or derived datasets. While data access agreements are legally binding if designed properly, their practical enforceability, especially across borders, is largely untested and remains a concern.¹¹ Especially where terms are associated with open access data, they are typically meant more as a means of communicating community norms to users.

19.3.3 *Monitoring of Data Use*

DACs may additionally develop tools and mechanisms to maintain ongoing oversight of downstream data uses. For instance, data users may be required to provide periodic reports regarding the projects in which data are being used. In addition, data users may be asked to report to the DAC the publications resulting from the data use, or issues arising from special conditions of access, e.g. risk management strategies for sensitive or potentially 'sensational' research, or return of incidental findings. Such oversight may enable the DACs to check compliance of the data uses, but implementation requires infrastructure and human resources that may be burdensome for DACs that do not have dedicated funding. There may also be important burdens – e.g. reporting or transparency obligations – placed on data users that discourage frivolous use. Research teams releasing data or DACs may have little ability to monitor data users or to directly sanction them for misuse, except by withdrawing or refusing access in the future. Some level of accountability is available via community reporting and

¹⁰ OECD, 'Recommendations on Health Data Governance', (OECD), www.oecd.org/els/health-systems/health-data-governance.htm.

¹¹ Burton et al., 'Policies and Strategies for Secondary Data Use of Data'; Global Alliance for Genomics & Health, 'GA4GH Accountability Policy 2016', (Global Alliance for Genomics & Health, 2016).

norms. Research institutions, funders, journals and databases themselves may have mechanisms to hold researchers accountable for respecting their commitments.¹²

19.3.4 *Maintaining Transparency*

The constitution of DACs shape how policies and governance mechanisms are implemented in practice. DACs are the site around which tensions between the competing interests of stakeholders may play out and therefore, examining how they do or do not maintain transparency allows scrutiny of those governance processes. DAC members may be part of the scientific team that generated the data, though the independence of members is often advocated in order to avoid conflicts of interest. Real or perceived conflicts of interest may arise where the researcher who collected the data restricts access to potential competitors, described as data ‘hugging’ or hoarding by those advocating data sharing.¹³ And yet, data producers have important expertise: they know the affordances and limits of the data as well as its provenance. In some DACs, this expertise is recognised by including members of the study team in an advisory role.¹⁴ Furthermore, all stakeholders should have some representation in governance of data access including as decision-making members of DACs. Stakeholder engagement may also comprise forms of transparency, for example through publication of high-quality plain language summaries to communicate how study data are, or will be, used.

19.4 BEST PRACTICE EXAMPLES

Depending on the organisation or its specific needs, data access governance can emphasise different governance-related values and goals.

19.4.1 *Multi-study Access: European Genome-Phenome Archive (EGA)*

An example of the local access management model is the collection of study DACs under the framework of the EGA. EGA is a database of all types of ‘sequence and genotype experiments, including case-control, population, and family studies, hosted at the European Bioinformatics Institute’.¹⁵ According to the EGA website: ‘The EGA will serve as a permanent archive that will archive several levels of data including the raw data (which could, for example, be re-analysed in the future by other algorithms) as well as the genotype calls provided by the submitters.’¹⁶ Data submitters via EGA maintain control over the downstream uses of datasets via DACs located in the original study or consortium. An advantage of local data access review is that data generators who are familiar with the dataset can stay involved in the process of review and inform the access review procedure. The disadvantage of this model is that the access control is entirely left to the local committees, making it hard if not impossible to track/audit whether all data access requests are being handled in a timely manner.

¹² Global Alliance for Genomics & Health, ‘GA4GH Accountability Policy 2016’.

¹³ Murtagh et al., ‘International Data Sharing’.

¹⁴ Murtagh et al., ‘METADAC Governance Infrastructure’.

¹⁵ European Genome-phenome Archive (EGA), ‘Introduction’, (EGA, 2019), www.ega-archive.org/about/introduction.

¹⁶ Ibid.

19.4.2 *Centralised Access: Database of Genotypes and Phenotypes (dbGaP)*

In contrast, dbGaP exemplifies a centralised approach to managing data access requests. The dbGaP is designed by the National Institutes of Health (NIH) to archive and distribute the results of studies that have investigated the interaction of genotype and phenotype. Within this database, sixteen DACs ‘review requests for consistency with any data use limitations and approve, disapprove or return requests for revision’, except for large studies in which a local DAC leads access review.¹⁷ The centralised access model seems advantageous for smaller research groups who lack resources to establish their own data access review infrastructure. However, the handling of data access requests centrally may lead to latency in data access, due to complex administrative arrangements.

19.4.3 *Tiered Access: International Cancer Genome Consortium/25K Initiative*

The International Cancer Genome Consortium (now called the 25K Initiative) was a large-scale genomics research initiative aiming to generate and share 25,000 whole genome sequences from fifteen jurisdictions to better understand the genetic changes occurring in different forms of cancer.¹⁸ The International Cancer Genome Consortium (ICGC) adopted a tiered access approach, with open access for data unlikely to be linked to other data that could re-identify individual participants, and controlled access for more sensitive data such as raw sequence and genotype files – though the exact data types in these two categories evolved over time.¹⁹ These more sensitive data can only be accessed through the Data Access Committee Office (DACO) to protect the privacy and reasonable expectations of study participants, uphold scientific community norms of attribution and publication priority, and ensure the impartiality of access decisions. The DACO reviews the purpose and relevance of research proposals, and the trustworthiness of applicants to protect participant privacy and data security. The ICGC adopted a plain language access agreement restricting users from establishing parasitic intellectual property on primary data or attempting to re-identify individual participants, with signatures from the principal investigator and institutional signing official. Recognising that requirements for ethics review vary from country to country, the DACO asks applicants to indicate if their study of ICGC data requires local ethics approval.

19.4.4 *Independent, Interdisciplinary Access Involving Stakeholder Participation in Decisions: METADAC (Managing Ethical, Socio-Technical and Administrative Issues in Data Access)*²⁰

METADAC provides data access governance for only the most sensitive data and data combinations (as well as sample access). While separating access in this way produces a complex data governance setting for researchers, the devolvement to different degrees of scrutiny for differently risky data allows resources for human-mediated decision making, where this is necessary and

¹⁷ D. Paltoo et al., ‘Data Use Under the NIH GWAS Data Sharing Policy and Future Directions’, (2014) *Nature Genetics*, 46(9), 934–938, 934.

¹⁸ International Cancer Genome Consortium, ‘About Us’, (International Cancer Genome Consortium, 2018), www.icgc.org/about-us.

¹⁹ Y. Joly et al., ‘Data Sharing in the Post-Genomic World: The Experience of the International Cancer Genome Consortium (ICGC) Data Access Compliance Office (DACO)’, (2012) *PLoS Computational Biology*, 8(7), e1002549.

²⁰ Murtagh et al., ‘METADAC Governance Infrastructure’.

allows administrative or algorithm-based decisions for low risk data types. The human-mediated decisions made by METADAC include a proportionate review process for routine-but-sensitive data access applications and full committee decision-making for the remaining sensitive data access applications. The METADAC committee comprises a highly multidisciplinary committee, including study-facing members (currently drawn from the participants of longitudinal studies not regulated by METADAC), with non-voting representation from the studies (including their technical teams) and the funders of these studies. Data access under METADAC does not require additional ethical approval as data sharing is based on tissue bank approval under the Human Tissue Act 2004,²¹ study ethical approval and/or explicit participant consent to sharing. METADAC's key criteria for access follow precisely the questions outlined in '*Respect for the provisions of ethical agreements*' above. The METADAC committee does not review the scientific merit of data access applications except in the case of finite resources (i.e. samples).

19.4.5 Data Producers' Rights and Interests: ClinicalStudyDataRequest

ClinicalStudyDataRequest.com is a portal facilitating access to patient-level data from clinical studies carried out by pharmaceutical companies and academic researchers.²² The portal involves independent review of proposals as well as protections for participant privacy and confidentiality. A major differentiator of this access model from the publicly funded genomic research context is protection of commercial interests. For pharmaceutical company-sponsored trials, the data sharing agreement requires users to keep all information provided confidential, in part to protect commercially sensitive information.²³ The user must also agree to give the sponsor an exclusive licence to any new intellectual property generated from the study. The agreement also requires users to publish or otherwise publicly disclose their results, which helps to ensure research is pursued for verification rather than commercial purposes.

19.4.6 Transparency and Reflexive Governance: UK Biobank (Ethics and Governance Framework)

In the late 2000s, in what would be an example of reflexive data access governance,²⁴ the UK Biobank revised its Ethics and Governance Framework (to address challenges that were current at the time). More specifically, the UK Biobank had originally committed to destroy the data of participants who chose to withdraw from the biobank. However, it soon realised that it could not uphold this commitment due to technical issues.²⁵ These issues included the establishment of IT systems that made it impossible to destroy data completely in order 'to protect the integrity and security of those people who have taken part'.²⁶ One year after identifying these issues, the

²¹ This Act applies to England, Northern Ireland and Wales. With the exception of section 45, which regulates DNA analysis, the Act does not extend to Scotland.

²² ClinicalStudyDataRequest, 'Home', (ClinicalStudyDataRequest), www.clinicalstudydatarequest.com/Default.aspx.

²³ ClinicalStudyDataRequest, 'Data Sharing Agreement', (ClinicalStudyDataRequest), www.clinicalstudydatarequest.com/Help/Help-Data-Sharing-Agreement.aspx.

²⁴ G. Laurie, 'Reflexive Governance in Biobanking: On the Value of Policy Led Approaches and the Need to Recognise the Limits of Law', (2011) *Human Genetics*, 130(3), 347–356.

²⁵ UK Biobank, 'Revision of the UK Biobank Ethics and Governance Framework: 'No Further Use' Withdrawal Option', (UK Biobank, 2007), www.egcukbiobank.org.uk/sites/default/files/Right%20to%20withdraw%20from%20UK%20Biobank.pdf.

²⁶ UK Biobank, "'NO FURTHER USE" Withdrawal Option – UK Biobank's Commitment to Your Wishes', (UK Biobank, 2011), www.ukbiobank.ac.uk/.

UK Biobank discussed and agreed with its Ethics and Governance Council to amend the scope of its commitment: rather than destroying participant data, the biobank would commit to ensure these data would be made completely unusable. UK Biobank subsequently revised both the participant information materials and governance frameworks not only to reflect this change, but to also describe the underlying reasons. In effect, such transparency and reflexivity could increase participant trust, and ultimately, participation in biobanks.

19.5 CHALLENGES AND FUTURE DIRECTIONS

19.5.1 *Resources, Effectiveness and Efficiency of Data Access Governance*

Not all research teams or repositories have the guidance, resources or expertise to establish responsible data access governance. Adequate support from funding agencies and institutions is key. This support may include establishing community data repositories to store and manage access on behalf of researchers.

Concerns regarding the workload of DACs in manually reviewing data access requests are the basis for emerging innovations around automation of at least some parts of the data access review.²⁷ One example of such efforts has been to automate the review of the conformity of the proposed data use with any use restrictions attached to the dataset – e.g. a consent agreement restricting use to non-commercial or disease specific research. In this regard, a recent initiative supported by the Global Alliance for Genomics and Health (GA4GH) developed a matrix for machine-readable consent forms. While these technical approaches will support the work substantially, there will likely always be a need for human review of the most sensitive or disclosive data access requests.

19.5.2 *Coordination between Oversight Bodies*

Oversight of access to biomedical databases would benefit considerably from further coordination between the relevant oversight bodies, such as DACs and RECs.²⁸ A single data-intensive research project may require access to multiple resources governed by multiple DACs, meaning multiple forms, reviews and delays. Multi-study DACs, such as METADAC, address the problem of repeated and time-consuming access processes. Requirements for multiple approvals from both ethics committees and DACs are dealt with in different ways. In the UK, for example, ethics review under the Human Tissue Act 2004 provides for broad approval for data sharing at the biobank level if relevant consents and other ethical safeguards are in place; permission for specific data access requests then only needs approval from the relevant DAC. Where national legislation is not in place, local or consortia arrangements are possible. The ICGC have disentangled ethics review from data access request review. Indeed, the ICGC's DACO consistently maintains that its DAC is not an ethics review committee and that it should not evaluate the consent forms of users or their

²⁷ S. Dyke et al., 'Registered Access: Authorizing Data Access', (2018) *European Journal of Human Genetics*, 26(12), 1721–1731.

²⁸ E. Dove et al., 'Ethics Review for International Data-intensive Research', (2016) *Science*, 351(6280), 1399–1400; M. Shabani et al., 'Oversight of Genomic Data Sharing: What Roles for Ethics and Data Access Committees?', (2017) *Biopreservation and Biobanking*, 15(5), 469–474.

research protocols, relying instead ‘on the local ethics processes of the data users without imposing another layer of ethics review requirements on them’.²⁹

19.5.3 *Harmonisation of Access Policies and Processes*

Interoperability of data access governance supports an important goal of data science, which is to combine similar datasets together to increase statistical power and thereby produce greater scientific insight. Access arrangements are currently fragmented, differing across countries, institutions and databases. These fragmented access arrangements have the potential to undermine usability of databases and produce data silos as users battle to conform to a variety of – sometimes contradictory – access requirements and conditions. Undertaking multiple roughly similar access processes to access different databases is not only burdensome, it also does not necessarily improve participant/data subject protections. Different aspects of access review can be streamlined so that they do not have to be repeated every time a researcher seeks access. Interoperability and predictability can be improved where different data stewards adopt standard access criteria. Central access portals could accept single requests to multiple data resources. This may be possible even where there are differences between the access conditions applying to the datasets. A step further would be to delegate certain aspects of access review. A common authentication body, for example, could be responsible for establishing the identity and affiliation of researchers, who could then present a single set of credentials to different access bodies.³⁰

19.6 CONCLUSION

Data access governance has an ultimate goal of taking into account and maintaining balance between the rights and interests of various stakeholders involved in data sharing. A central aim of data access governance, of course, is to promote broad access to data to advance knowledge and improve human health. In doing so, it is essential to have a comprehensive overview of the rights and interests of the involved parties that might be in contrast with each other when establishing rules for data access reviews and approvals.

In view of increasing data sharing among researchers, it is crucial to ensure the DACs and RECs have sufficient resources to achieve the ultimate goals of access review, namely transparency, fairness and proportionality. In doing so, adopting a number of already proposed approaches would be advantageous, including – partly – automating the process of access review and introducing light-touch forms of review when sharing non-sensitive data.

Technological advancements could lead to heightened risks of re-identification of individuals when sharing sensitive health related data. Therefore, it is important to ensure the adopted governance mechanisms include adequate safeguards when sharing data. In addition, in establishing governance mechanisms, attention should be paid to the social values underpinning data sharing. Thereby, the focus of data governance should not be limited to only protecting the individual rights and interests of the involved parties, but also to fostering social values that can arise from promoting responsible data sharing.

²⁹ Joly et al., ‘Data Sharing in the Post-Genomic World’.

³⁰ Global Alliance for Genomics & Health, ‘GA4GH Data Use and Researcher ID Work Stream’, (GA4GH), <https://ga4gh-duri.github.io/>.